RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/539,560
Source: PCT
Date Processed by STIC: 01/31/2006

ENTERED

PCT

RAW SEQUENCE LISTING DATE: 01/31/2006
PATENT APPLICATION: **US/10/539,560** TIME: 16:14:19

Input Set : **A:\18765218.APP**
Output Set: **N:\CRF4\01312006\J539560.raw**

```
  3 <110> APPLICANT: YAMAKI, TOSHIFUMI
  4       BANBA, SHINICHI
  5       MATOISHI, KAORI
  6       ITO, KYOSHI
  7       KOBAYASHI, HIDEKI
  8       TANAKA, EISHI
  9       OIKAWA, TOSHIHIRO
 11 <120> TITLE OF INVENTION: NOVEL NITRILE HYDRATASE
 13 <130> FILE REFERENCE: 018765-218
 15 <140> CURRENT APPLICATION NUMBER: 10/539,560
C--> 16 <141> CURRENT FILING DATE: 2005-06-17
 18 <150> PRIOR APPLICATION NUMBER: PCT/JP03/016014
 19 <151> PRIOR FILING DATE: 2003-12-15
 21 <150> PRIOR APPLICATION NUMBER: JP 2003-379280
 22 <151> PRIOR FILING DATE: 2003-11-10
 24 <150> PRIOR APPLICATION NUMBER: JP 2002-368360
 25 <151> PRIOR FILING DATE: 2002-12-19
 27 <160> NUMBER OF SEQ ID NOS: 142
 29 <170> SOFTWARE: PatentIn Ver. 3.3
 31 <210> SEQ ID NO: 1
 32 <211> LENGTH: 205
 33 <212> TYPE: PRT
 34 <213> ORGANISM: Pseudonocardia thermophila
 36 <400> SEQUENCE: 1
 37 Met Thr Glu Asn Ile Leu Arg Lys Ser Asp Glu Glu Ile Gln Lys Glu
 38   1               5                  10                  15
 40 Ile Thr Ala Arg Val Lys Ala Leu Glu Ser Met Leu Ile Glu Gln Gly
 41              20                  25                  30
 43 Ile Leu Thr Thr Ser Met Ile Asp Arg Met Ala Glu Ile Tyr Glu Asn
 44          35                  40                  45
 46 Glu Val Gly Pro His Leu Gly Ala Lys Val Val Val Lys Ala Trp Thr
 47      50                  55                  60
 49 Asp Pro Glu Phe Lys Lys Arg Leu Leu Ala Asp Gly Thr Glu Ala Cys
 50  65                  70                  75                  80
 52 Lys Glu Leu Gly Ile Gly Gly Leu Gln Gly Glu Asp Met Met Trp Val
 53                  85                  90                  95
 55 Glu Asn Thr Asp Glu Val His His Val Val Val Cys Thr Leu Cys Ser
 56              100                 105                 110
 58 Cys Tyr Pro Trp Pro Val Leu Gly Leu Pro Pro Asn Trp Phe Lys Glu
 59          115                 120                 125
 61 Pro Gln Tyr Arg Ser Arg Val Val Arg Glu Pro Arg Gln Leu Leu Lys
 62      130                 135                 140
 64 Glu Glu Phe Gly Phe Glu Val Pro Pro Ser Lys Glu Ile Lys Val Trp
```

cpg-6)

RAW SEQUENCE LISTING DATE: 01/31/2006
PATENT APPLICATION: US/10/539,560 TIME: 16:14:19

Input Set : A:\18765218.APP
Output Set: N:\CRF4\01312006\J539560.raw

```
65 145                 150                 155                 160
67 Asp Ser Ser Ser Glu Met Arg Phe Val Val Leu Pro Gln Arg Pro Ala
68                  165                 170                 175
70 Gly Thr Asp Gly Trp Ser Glu Glu Glu Leu Ala Thr Leu Val Thr Arg
71              180                 185                 190
73 Glu Ser Met Ile Gly Val Glu Pro Ala Lys Ala Val Ala
74          195                 200                 205
77 <210> SEQ ID NO: 2
78 <211> LENGTH: 233
79 <212> TYPE: PRT
80 <213> ORGANISM: Pseudonocardia thermophila
82 <400> SEQUENCE: 2
83 Met Asn Gly Val Tyr Asp Val Gly Gly Thr Asp Gly Leu Gly Pro Ile
84   1               5                  10                  15
86 Asn Arg Pro Ala Asp Glu Pro Val Phe Arg Ala Glu Trp Glu Lys Val
87              20                  25                  30
89 Ala Phe Ala Met Phe Pro Ala Thr Phe Arg Ala Gly Phe Met Gly Leu
90          35                  40                  45
92 Asp Glu Phe Arg Phe Gly Ile Glu Gln Met Asn Pro Ala Glu Tyr Leu
93      50                  55                  60
95 Glu Ser Pro Tyr Tyr Trp His Trp Ile Arg Thr Tyr Ile His His Gly
96  65                  70                  75                  80
98 Val Arg Thr Gly Lys Ile Asp Leu Glu Glu Leu Glu Arg Arg Thr Gln
99                  85                  90                  95
101 Tyr Tyr Arg Glu Asn Pro Asp Ala Pro Leu Pro Glu His Glu Gln Lys
102             100                 105                 110
104 Pro Glu Leu Ile Glu Phe Val Asn Gln Ala Val Tyr Gly Gly Leu Pro
105         115                 120                 125
107 Ala Ser Arg Glu Val Asp Arg Pro Pro Lys Phe Lys Glu Gly Asp Val
108     130                 135                 140
110 Val Arg Phe Ser Thr Ala Ser Pro Lys Gly His Ala Arg Arg Ala Arg
111 145                 150                 155                 160
113 Tyr Val Arg Gly Lys Thr Gly Thr Val Val Lys His His Gly Ala Tyr
114                 165                 170                 175
116 Ile Tyr Pro Asp Thr Ala Gly Asn Gly Leu Gly Glu Cys Pro Glu His
117             180                 185                 190
119 Leu Tyr Thr Val Arg Phe Thr Ala Gln Glu Leu Trp Gly Pro Glu Gly
120         195                 200                 205
122 Asp Pro Asn Ser Ser Val Tyr Tyr Asp Cys Trp Glu Pro Tyr Ile Glu
123     210                 215                 220
125 Leu Val Asp Thr Lys Ala Ala Ala Ala
126 225                 230
129 <210> SEQ ID NO: 3
130 <211> LENGTH: 618
131 <212> TYPE: DNA
132 <213> ORGANISM: Pseudonocardia thermophila
134 <400> SEQUENCE: 3
135 atgaccgaga acatcctgcg caagtcggac gaggagatcc agaaggagat cacggcgcgg 60
136 gtcaaggccc tggagtcgat gctcatcgaa cagggcatcc tcaccacgtc gatgatcgac 120
```

RAW SEQUENCE LISTING DATE: 01/31/2006
PATENT APPLICATION: US/10/539,560 TIME: 16:14:19

Input Set : A:\18765218.APP
Output Set: N:\CRF4\01312006\J539560.raw

cggatggccg agatctacga gaacgaggtc ggcccgcacc tcggcgcgaa ggtcgtcgtg 180
aaggcctgga ccgacccgga gttcaagaag cgtctgctcg ccgacggcac cgaggcctgc 240
aaggagctcg gcatcggcgg cctgcagggc gaggacatga tgtgggtgga gaacaccgac 300
gaggtccacc acgtcgtcgt gtgcacgctc tgctcctgct acccgtggcc ggtgctgggg 360
ctgccgccga actggttcaa ggagccgcag taccgctccc gcgtggtgcg tgagccccgg 420
cagctgctca aggaggagtt cggcttcgag gtcccgccga gcaaggagat caaggtctgg 480
gactccagct ccgagatgcg cttcgtcgtc ctcccgcagc gccccgcggg caccgacggg 540
tggagcgagg aggagctcgc caccctcgtc acccgcgagt cgatgatcgg cgtcgaaccg 600
gcgaaggcgg tcgcgtga 618

<210> SEQ ID NO: 4
<211> LENGTH: 702
<212> TYPE: DNA
<213> ORGANISM: Pseudonocardia thermophila

<400> SEQUENCE: 4
atgaacggcg tgtacgacgt cggcggcacc gatgggctgg gcccgatcaa ccggcccgcg 60
gacgaaccgg tcttccgcgc cgagtgggag aaggtcgcgt tcgcgatgtt cccggcgacg 120
ttccgggccg gcttcatggg cctggacgag ttccggttcg gcatcgagca gatgaacccg 180
gccgagtacc tcgagtcgcc gtactactgg cactggatcc gcacctacat ccaccacggc 240
gtccgcaccg gcaagatcga tctcgaggag ctggagcgcc gcacgcagta ctaccgggag 300
aaccccgacg ccccgctgcc cgagcacgag cagaagccgg agttgatcga gttcgtcaac 360
caggccgtct acggcgggct gcccgcaagc cgggaggtcg accgaccgcc caagttcaag 420
gagggcgacg tggtgcggtt ctccaccgcg agcccgaagg gccacgcccg gcgcgcgcgg 480
tacgtgcgcg gcaagaccgg gacggtggtc aagcaccacg gcgcgtacat ctacccggac 540
accgccggca acggcctggg cgagtgcccc gagcacctct acaccgtccg cttcacggcc 600
caggagctgt gggggccgga aggggacccg aactccagcg tctactacga ctgctgggag 660
ccctacatcg agctcgtcga cacgaaggcg gccgcggcat ga 702

<210> SEQ ID NO: 5
<211> LENGTH: 144
<212> TYPE: PRT
<213> ORGANISM: Pseudonocardia thermophila

<400> SEQUENCE: 5
Met Ser Ala Glu Ala Lys Val Arg Leu Lys His Cys Pro Thr Ala Glu
1 5 10 15
Asp Arg Ala Ala Ala Asp Ala Leu Leu Ala Gln Leu Pro Gly Gly Asp
20 25 30
Arg Ala Leu Asp Arg Gly Phe Asp Glu Pro Trp Gln Leu Arg Ala Phe
35 40 45
Ala Leu Ala Val Ala Ala Cys Arg Ala Gly Arg Phe Glu Trp Lys Gln
50 55 60
Leu Gln Gln Ala Leu Ile Ser Ser Ile Gly Glu Trp Glu Arg Thr His
65 70 75 80
Asp Leu Asp Asp Pro Ser Trp Ser Tyr Tyr Glu His Phe Val Ala Ala
85 90 95
Leu Glu Ser Val Leu Gly Glu Glu Gly Ile Val Glu Pro Glu Ala Leu
100 105 110
Asp Glu Arg Thr Ala Glu Val Leu Ala Asn Pro Pro Asn Lys Asp His
115 120 125
His Gly Pro His Leu Glu Pro Val Ala Val His Pro Ala Val Arg Ser
130 135 140

RAW SEQUENCE LISTING DATE: 01/31/2006
PATENT APPLICATION: **US/10/539,560** TIME: 16:14:19

Input Set : **A:\18765218.APP**
Output Set: **N:\CRF4\01312006\J539560.raw**

205 <210> SEQ ID NO: 6
206 <211> LENGTH: 435
207 <212> TYPE: DNA
208 <213> ORGANISM: Pseudonocardia thermophila
210 <400> SEQUENCE: 6
211 gtgagcgccg aggcgaaggt ccgcctgaag cactgcccca cggccgagga ccgggcggcg 60
212 gccgacgcgc tgctcgcgca gctgcccggc ggcgaccgcg cgctcgaccg cggcttcgac 120
213 gagccgtggc agctgcgggc gttcgcgctg gcggtcgcgg cgtgcagggc gggccggttc 180
214 gagtggaagc agctgcagca ggcgctgatc tcctcgatcg gggagtggga gcgcacccac 240
215 gatctcgacg atccgagctg gtcctactac gagcacttcg tcgccgcgct ggaatccgtg 300
216 ctcggcgagg aagggatcgt cgagccggag gcgctggacg agcgcaccgc ggaggtcttg 360
217 gccaacccgc cgaacaagga tcaccatgga ccgcatctgg agcccgtcgc ggtccacccg 420
218 gccgtgcggt cctga 435
221 <210> SEQ ID NO: 7
222 <211> LENGTH: 18
223 <212> TYPE: DNA
224 <213> ORGANISM: Artificial Sequence
226 <220> FEATURE:
227 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
228 primer
230 <400> SEQUENCE: 7
231 aacatcatgc gcaagtcg 18
234 <210> SEQ ID NO: 8
235 <211> LENGTH: 17
236 <212> TYPE: DNA
237 <213> ORGANISM: Artificial Sequence
239 <220> FEATURE:
240 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
241 primer
243 <400> SEQUENCE: 8
244 gttttcccag tcacgac 17
247 <210> SEQ ID NO: 9
248 <211> LENGTH: 20
249 <212> TYPE: DNA
250 <213> ORGANISM: Artificial Sequence
252 <220> FEATURE:
253 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
254 primer
256 <400> SEQUENCE: 9
257 ggccagtgcc tagcttacat 20
260 <210> SEQ ID NO: 10
261 <211> LENGTH: 17
262 <212> TYPE: DNA
263 <213> ORGANISM: Artificial Sequence
265 <220> FEATURE:
266 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
267 primer
269 <400> SEQUENCE: 10
270 caggaaacag ctatgac 17

RAW SEQUENCE LISTING DATE: 01/31/2006
PATENT APPLICATION: US/10/539,560 TIME: 16:14:19

Input Set : A:\18765218.APP
Output Set: N:\CRF4\01312006\J539560.raw

273 <210> SEQ ID NO: 11
274 <211> LENGTH: 18
275 <212> TYPE: DNA
276 <213> ORGANISM: Artificial Sequence
278 <220> FEATURE:
279 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
280 primer
282 <400> SEQUENCE: 11
283 aacatcacgc gcaagtcg 18
286 <210> SEQ ID NO: 12
287 <211> LENGTH: 18
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
293 primer
295 <400> SEQUENCE: 12
296 aacatcgcgc gcaagtcg 18
299 <210> SEQ ID NO: 13
300 <211> LENGTH: 18
301 <212> TYPE: DNA
302 <213> ORGANISM: Artificial Sequence
304 <220> FEATURE:
305 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
306 primer
308 <400> SEQUENCE: 13
309 aacatcgtgc gcaagtcg 18
312 <210> SEQ ID NO: 14
313 <211> LENGTH: 18
314 <212> TYPE: DNA
315 <213> ORGANISM: Artificial Sequence
317 <220> FEATURE:
318 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
319 primer
321 <400> SEQUENCE: 14
322 atcacggtgc gggtcaag 18
325 <210> SEQ ID NO: 15
326 <211> LENGTH: 18
327 <212> TYPE: DNA
328 <213> ORGANISM: Artificial Sequence
330 <220> FEATURE:
331 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
332 primer
334 <400> SEQUENCE: 15
335 acgtcgttga tcgaccgg 18
338 <210> SEQ ID NO: 16
339 <211> LENGTH: 18
340 <212> TYPE: DNA
341 <213> ORGANISM: Artificial Sequence

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 01/31/2006
PATENT APPLICATION: **US/10/539,560** TIME: 16:14:20

Input Set : **A:\18765218.APP**
Output Set: **N:\CRF4\01312006\J539560.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:110; N Pos. 14,15,16
Seq#:111; N Pos. 7,8,9
Seq#:112; N Pos. 7,8,9
Seq#:113; N Pos. 7,8,9
Seq#:114; N Pos. 7,8,9
Seq#:115; N Pos. 7,8,9
Seq#:116; N Pos. 7,8,9
Seq#:117; N Pos. 7,8,9
Seq#:118; N Pos. 7,8,9
Seq#:119; N Pos. 7,8,9
Seq#:120; N Pos. 7,8,9
Seq#:121; N Pos. 7,8,9
Seq#:122; N Pos. 7,8,9
Seq#:123; N Pos. 7,8,9
Seq#:124; N Pos. 7,8,9
Seq#:125; N Pos. 7,8,9
Seq#:126; N Pos. 7,8,9
Seq#:127; N Pos. 7,8,9
Seq#:128; N Pos. 7,8,9
Seq#:129; N Pos. 7,8,9
Seq#:130; N Pos. 7,8,9
Seq#:131; N Pos. 7,8,9
Seq#:132; N Pos. 7,8,9
Seq#:133; N Pos. 7,8,9
Seq#:134; N Pos. 7,8,9
Seq#:135; N Pos. 7,8,9
Seq#:136; N Pos. 7,8,9
Seq#:137; N Pos. 7,8,9
Seq#:138; N Pos. 7,8,9
Seq#:139; N Pos. 7,8,9
Seq#:142; Xaa Pos. 1,3,8,9,10,11

VERIFICATION SUMMARY DATE: 01/31/2006
PATENT APPLICATION: **US/10/539,560** TIME: 16:14:20

Input Set : **A:\18765218.APP**
Output Set: **N:\CRF4\01312006\J539560.raw**

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1666 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:104
L:1802 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:110 after pos.:0
L:1820 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:111 after pos.:0
L:1838 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:112 after pos.:0
L:1856 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:113 after pos.:0
L:1874 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:114 after pos.:0
L:1892 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:115 after pos.:0
L:1910 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:116 after pos.:0
L:1928 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:117 after pos.:0
L:1946 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118 after pos.:0
L:1964 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:119 after pos.:0
L:1982 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:120 after pos.:0
L:2000 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:121 after pos.:0
L:2018 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:122 after pos.:0
L:2036 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123 after pos.:0
L:2054 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:124 after pos.:0
L:2072 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:125 after pos.:0
L:2090 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:126 after pos.:0
L:2108 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:127 after pos.:0
L:2126 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:128 after pos.:0
L:2144 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:129 after pos.:0
L:2162 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:130 after pos.:0
L:2180 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:131 after pos.:0
L:2198 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:132 after pos.:0
L:2216 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:133 after pos.:0
L:2234 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:134 after pos.:0
L:2252 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:135 after pos.:0
L:2270 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:136 after pos.:0
L:2288 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:137 after pos.:0
L:2306 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:138 after pos.:0
L:2324 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139 after pos.:0
L:2466 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:142 after pos.:0